



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2010

Welche Arten von Studiendesigns gibt es und wie werden sie korrekt eingesetzt?

Held, U

DOI: <https://doi.org/10.4414/smf.2010.07304>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-46204>

Journal Article

Published Version

Originally published at:

Held, U (2010). Welche Arten von Studiendesigns gibt es und wie werden sie korrekt eingesetzt? Swiss Medical Forum, 10(41):712-714.

DOI: <https://doi.org/10.4414/smf.2010.07304>

Welche Arten von Studiendesigns gibt es und wie werden sie korrekt eingesetzt?

Ulrike Held

Horten-Zentrum, UniversitätsSpital, Zürich

Im letzten Artikel dieser Serie geht es um Studientypen bzw. Studiendesigns, die zur Beantwortung von Fragen in der medizinischen Forschung dienen. Das Ziel einer solchen Studie ist es, eine Antwort auf die primäre Fragestellung geben zu können, also eine Evidenz für oder gegen z.B. das Vorliegen eines bestimmten Behandlungseffekts zu finden. Ob dies am Ende der Studie möglich sein wird, hängt von vielen Faktoren wie z.B. dem Stichprobenumfang, der geeigneten Auswahl der Studienteilnehmer, der Randomisierung, der Variation innerhalb der Daten, der gewählten Analyseverfahren ... und eben von der richtigen Wahl des Studiendesigns ab. Die gängigen epidemiologischen Studiendesigns umfassen Kohortenstudien, Querschnittsstudien und Fall-Kontroll-Studien, in Untersuchungen mit experimentellem Charakter verwendet man randomisierte, kontrollierte Studien, Cross-over-Studien und Äquivalenzstudien, und schliesslich gehen wir noch kurz auf Metaanalysen ein.

Kohortenstudien

Kohortenstudien sind Longitudinalstudien: Eine grosse Zahl von Studienteilnehmern, die die Kohorte bilden, wird prospektiv über einen gewissen Zeitraum (Follow-up) oder bis ein definiertes Ereignis eintritt, beobachtet. Dieses Ereignis kann z.B. Herzinfarkt, Auftreten von Diabetes mellitus oder Krebs sein. Dabei werden Daten über wichtige Einflussgrössen, vor allem aber die Exposition des untersuchten Risikofaktors über die Zeit, gesammelt und in Bezug auf die Zielgrösse ausgewertet. Man bezeichnet Kohortenstudien auch als Follow-up-Studien.

Kohortenstudien eignen sich dafür, kausale Zusammenhänge aufzudecken, da die Exposition zeitlich vor dem Eintreten des Ereignisses liegt und der Expositionsstatus bei allen Individuen bekannt ist. Ein weiterer Vorteil von Kohortenstudien liegt in der zeitnahen Erfassung des Expositionsstatus, der mitunter zeitlich variieren kann, weil somit eine Expositions-Wirkungs-Beziehung sehr präzise geschätzt werden kann. Allerdings sind Kohortenstudien sehr zeit- und kostenintensiv, vor allem bei Krankheiten mit einer langen Latenzzeit oder bei sehr seltenen Krankheiten. Wenn nun die Follow-up-Zeit sehr lang ist, ergeben sich weitere Probleme wie z.B. die Vergleichbarkeit der Untersuchungsbedingungen zu Beginn und Ende des Follow-up, die Einführung neuer Diagnosetechniken oder nichtzufällige Ausfälle von Mitgliedern der Studienpopulation. Dies stellt ein Problem dar, sobald die Ausfälle vom Untersuchungsziel abhängen. Weiterhin muss sichergestellt sein, dass alle

wichtigen Einflussfaktoren von Beginn der Studie an erfasst wurden, ein nachträgliches Hinzunehmen weiterer Faktoren ist sehr schwierig. Also sind Kohortenstudien prinzipiell geeignet, ätiologische Fragestellungen zu beantworten, andererseits kommt es in der praktischen Durchführung häufig zu gravierenden Problemen.

Querschnittsstudien

Querschnittsstudien (cross-sectional studies) werden häufig in der epidemiologischen Forschung verwendet. Mit ihrer Hilfe kann die Inzidenz oder Prävalenz einer Erkrankung innerhalb einer Population geschätzt werden. Die Daten können relativ schnell erhoben werden, weil kein Follow-up nötig ist. Exposition und Wirkung werden zum gleichen Zeitpunkt oder innerhalb eines kurzen Zeitfensters erfasst. Da häufig ein langer Zeitraum zwischen Exposition und Krankheitsausbruch liegt, sollten Querschnittsstudien nicht zum Nachweis einer Kausalbeziehung verwendet werden, sondern eher deskriptiv zur Beschreibung des Ist-Zustands oder zur Generierung neuer Hypothesen. Wichtig ist hierbei, auf eine repräsentative Auswahl der Studienteilnehmer zu achten.

Vorteile von Querschnittsstudien sind ihre kurze Dauer, die mit einem geringen Kostenaufwand verbunden ist. Sind die untersuchten Krankheiten von langer Dauer und nicht zu selten und die untersuchte Exposition eine Dauergewohnheit, wie z.B. Rauchen oder Wohnen, kann dieser Studientyp durchaus angewendet werden. Nachteile von Querschnittsstudien sind die geringe Präzision in der Erfassung der Exposition und ihr zeitlicher Zusammenhang mit dem Auftreten der Krankheit. Bei seltenen Krankheiten oder Krankheiten von kurzer Dauer sind Querschnittsstudien nicht geeignet. Weiterhin kann die Qualität einer Querschnittsstudie darunter leiden, dass eine bestimmte Art von Auswahlverzerrung die Ergebnisse beeinflusst. Dies nämlich, wenn sich durch das Ausbrechen der untersuchten Krankheit der Expositionsstatus verändert, wie z.B. bei Studien über Kinder mit Atemwegsproblemen, in denen Familien aus schadstoffbelasteten Gebieten häufig den Wohnort wechseln.

Fall-Kontroll-Studien

Fall-Kontroll-Studien (case-control studies) eignen sich zur Untersuchung von seltenen Krankheiten oder Ereignissen. Hier können relativ einfach alle bekannten Fälle



Ulrike Held

Die Autoren erklären, dass sie keine Interessenkonflikte im Zusammenhang mit diesem Beitrag haben.

(cases) gesammelt und retrospektiv die relevanten Einflussgrössen dazu erhoben werden. Etwas schwieriger ist dann die Auswahl von passenden Kontrollen (controls). Diese werden häufig über das sog. Matching gefunden. Beim Matching oder der Paarbildung möchte man die Gruppe der Fälle und die Gruppe der Kontrollen in Bezug auf wichtige Störgrössen homogenisieren. Weiss man z.B., dass Alter und Geschlecht einen Einfluss auf die untersuchte Krankheit haben, so sucht man sich zu den Fällen (mit bekanntem Alter und Geschlecht) entsprechende Kontrollen, die das gleiche Alter und Geschlecht haben. Alter und Geschlecht wären hier also Matching-Variablen. Man unterscheidet zwischen individuellem Matching (1:1) und dem sog. Häufigkeits-matching (m:n), bei dem nur die Verteilung der Matching-Variablen in beiden Gruppen vergleichbar sein muss. Hat man nun die Störgrössen durch Homogenisieren in den Vergleichsgruppen «ausgeschaltet», kann man die Fälle und Kontrollen in Bezug auf andere Risikofaktoren, wie z.B. Raucherstatus oder Ernährungsgewohnheiten, vergleichen. Allerdings ist die Betrachtung der zeitlichen Abfolge von Ursache und Wirkung retrospektiv und die Etablierung einer ätiologischen Beziehung zwischen Exposition und Krankheit indirekt, indem alle anderen bekannten Störgrössen ausgeschlossen werden.

Fall-Kontroll-Studien sind ein weit verbreiteter Studientyp bei der Untersuchung von seltenen Erkrankungen, da diese Methode praktisch die einzige ist, die mit begrenztem Zeit- und Kostenaufwand noch verlässliche Ergebnisse liefern kann. Von Vorteil ist auch, dass bei Krankheiten mit langer Latenzzeit diese Zeit nicht im Laufe der Studie abgewartet werden muss, wie es bei den Kohortenstudien der Fall ist. Allerdings haben Fall-Kontroll-Studien auch Nachteile. Ein Kernproblem ist vor allem die Auswahl der Kontrollgruppe und die Sicherstellung, dass die Fall- und die Kontrollgruppe in punkto der wichtigsten Störgrössen homogen sind. Dies setzt insbesondere die Kenntnis aller wichtigen Störgrössen voraus. Ob dann auch tatsächlich eine homogene Verteilung dieser garantiert werden kann, ist zusätzlich fraglich. Selbst für den Fall, dass wirklich vergleichbare Gruppen von Fällen und Kontrollen vorliegen, kann die korrekte Analyse dadurch erschwert werden, dass der Forscher die Wirkung der bekannten Risikofaktoren «nur noch» nachweisen muss und somit nicht unvoreingenommen an die Auswertung herangehen kann. Weiterhin ist bei Fall-Kontroll-Studien die Expositionsbestimmung retrospektiv schwierig bzw. ungenau. Dies kann vor allem zu Verzerrungen führen, wenn den Studienteilnehmern die zu untersuchenden Risikofaktoren bekannt sind. So kann man davon ausgehen, dass z.B. Eltern von Kindern mit Fehlbildungen sich mehr Gedanken über mögliche Risikofaktoren vor oder während der Schwangerschaft gemacht haben als Eltern von gesunden Kindern, welche die mögliche Exposition im Laufe der Zeit vergessen haben.

Randomisierte kontrollierte Studien (RCT)

Die nachweislich beste Methode, um eine medizinische Fragestellung zu beantworten, ist die prospektive, randomisierte, plazebokontrollierte Studie. Die Randomi-

sierung, d.h. also die zufällige Aufteilung der Studienteilnehmer auf die Behandlungsgruppe(n) und die Vergleichsgruppe, ist hierbei von grosser Wichtigkeit, da durch sie die Vergleichbarkeit der Gruppen bezüglich aller Störgrössen erreicht werden soll, so dass Unterschiede in der Zielgrösse am Ende nur der Behandlung zugeschrieben werden können. Zusätzlich verhindert die Randomisierung auch das (möglicherweise unbewusste) nichtzufällige Zuteilen in Behandlungs- und Vergleichsgruppe durch den Arzt. Kontrolliert heisst die Studie, wenn die Ergebnisse der Behandlungs- oder Interventionsgruppe mit denen einer Vergleichs- oder Kontrollgruppe verglichen werden. Die Kontrollgruppe erhält entweder ein Scheinpräparat (Plazebo) oder die gängige Behandlungsmethode. Häufig sind randomisierte klinische Studien verblindet (maskiert): Ist die Studie einfach verblindet, dann kennt der Patient seine Zugehörigkeit zur Behandlungsgruppe nicht, und ist sie doppelt verblindet, dann weiss auch der Arzt nicht, in welche Gruppe der Patient randomisiert wurde. Die Verblindung soll verhindern, dass der Patient – und ggf. auch der Arzt – in seiner Bewertung des Krankheitsbildes von seinem Wissen über die Behandlung beeinflusst wird.

Cross-over-Studien

Cross-over-Studien sind klinische Studien, in denen jeder Proband nicht nur entweder in der Behandlungs- oder in der Kontrollgruppe ist, sondern alle Behandlungen nacheinander erhält. Die Randomisierung ist auch hier wichtig, aber sie bestimmt nur noch, in welcher Reihenfolge die Behandlungen erfolgen. Nach einer ersten erfolgten Behandlung kommt eine sog. Wash-out-Phase, um eine mögliche verzögerte Wirkung der Behandlung abzuwarten bzw. auszuschliessen. Danach sollte der Zielparameter sich wieder auf dem Niveau der Baseline vor Beginn der ersten Behandlungsphase befinden. Ist dies der Fall, kann mit der zweiten Behandlungsphase begonnen werden. Bei der Analyse von Daten aus Cross-over-Studien kann somit die Variabilität reduziert werden, denn der Effekt der Behandlung wird innerhalb jedes Patienten analysiert. Ein Vorteil von Cross-over-Studien ist also, dass der gleiche Effekt mit der gleichen Power auf der Basis einer geringeren Anzahl von Patienten nachgewiesen werden kann. Allerdings kommen natürlich Cross-over-Studien für gewisse medizinische Hypothesen nicht in Frage, wenn z.B. chirurgische Eingriffe mit alternativen Behandlungsmethoden verglichen werden sollen. Weitere Nachteile von Cross-over-Studien können auch die etwas längere Studiendauer sein und der sog. Carry-over-Effekt, falls nach Ende der Wash-out-Phase doch die erste Behandlung noch wirkt.

Äquivalenzstudien

Für die Zulassung von Generika werden Äquivalenzstudien verwendet. Deren Wirkung wird mit der Wirksamkeit der Originalpräparate verglichen, welche schon

zugelassen und auf dem Markt sind, und man möchte den Nachweis erbringen, dass das Generikum weder wesentlich schlechter noch wesentlich besser wirkt, dass es also äquivalent ist. Es wäre zu schwierig nachzuweisen, dass die Wirkung genau gleich ist. Die Bezeichnung Äquivalenz bezieht sich normalerweise auf die mittlere Wirksamkeit der Präparate. So könnte die Formulierung einer Äquivalenzhypothese wie folgt aussehen.

$$H_0: \mu_T \leq \mu_R - d \text{ oder } \mu_T \geq \mu_R + d$$

$$H_1: -d < \mu_T - \mu_R < d$$

wobei μ_R die mittlere Wirkung des Referenzpräparats und μ_T die mittlere Wirkung des Testpräparats oder Generikums ist. Typischerweise muss eine Abweichung in der Wirksamkeit um weniger als plus/minus 20% gezeigt werden, also $d = 0,2 \mu_R$ nachgewiesen werden. Wie man sieht, ist die Nullhypothese genau andersherum formuliert als bei der gewöhnlichen Untersuchung von wissenschaftlichen Fragestellungen.

Metaanalysen

Das Ziel von systematischen Reviews und Metaanalysen ist das Zusammenfassen von Primärstudien zu Metadaten und deren Auswertung mit entsprechenden statistischen Verfahren. Das Ergebnis einer Metaanalyse kann dann sein, eine medizinische Fragestellung unter Einbeziehung aller verfügbaren Informationen neu zu bewerten.

Als Effektgrößen, die gemeinsam analysiert werden, kommen alle statistischen Masse vor, hauptsächlich Odds Ratios, Risk Ratios, Risikodifferenz und Hazard Ratios. Allerdings gibt es wichtige Voraussetzungen, um überhaupt eine höhere Evidenz durch Metaanalysen zu erreichen: Die einbezogenen Primärstudien sollten sich bezüglich ihrer Studienpopulation möglichst ähnlich sein, und alle Studien sollten im besten Fall den gleichen primären Endpunkt haben.

Vorteile einer Metaanalyse sind einerseits, dass man alle Informationen zu einer relevanten Fragestellung gemeinsam analysieren kann und dafür auch anspruchsvolle statistische Methoden zur Verfügung stehen. Mit einer Metaanalyse kommt man schneller zu einem Ergebnis, als wenn man eine neue Primärstudie durchführen würde. Und basierend auf dem Ergebnis der Metaanalyse kann man sich ein besseres Bild machen, ob zu einer wissenschaftlichen Fragestellung eine erneute Primärstudie überhaupt noch nötig ist.

Allerdings gibt es auch viele Schwierigkeiten und Bias-Quellen bei Metaanalysen: Zunächst muss der Forscher alle relevanten Studien finden, im Prinzip sowohl die publizierten als auch die nichtpublizierten. Ansonsten droht eine Verzerrung der Metaanalyse: Man tendiert dazu, den tatsächlichen Effekt zu überschätzen (sog. Publikationsbias), weil Studien, die keinen Effekt zeigen, häufig nicht publiziert werden. Hinzu kommt, dass die einzelnen Studien häufig keine gute Qualität haben und dass kleine Studien eher grosse Effekte zeigen. Dies führt dann zu einer schlechten Datenqualität und zu verzerrten Ergebnissen der Metaanalyse.

Korrespondenz:

Dr. rer. nat. Ulrike Held
Horten-Zentrum
UniversitätsSpital Zürich
Postfach Nord
CH-8091 Zürich
ulrike.held@usz.ch

Empfohlene Literatur

- Bachmann L.M., Puhan M.A. und Steurer J. Patientenorientierte Forschung – Einführung in die Planung und Durchführung einer Studie. 2008. Huber Verlag, Bern.
- Held L., Ruffach C. und Seifert B. Einführung in die Biostatistik. 4. Auflage, Juli 2009. Abteilung Biostatistik, Institut für Sozial- und Präventivmedizin der Universität Zürich. <http://www.biostat.uzh.ch>.
- Hüsler J. und Zimmermann H. Statistische Prinzipien für medizinische Projekte. 4. Auflage 2006. Huber Verlag, Bern.
- Kreienbrock L. und Schach S. Epidemiologische Methoden. 4. Auflage 2005. Elsevier Verlag, München.